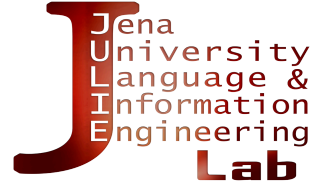




FRIEDRICH-SCHILLER-
UNIVERSITÄT
JENA



Quantifizierung von Emotionen in historischer Sprache

Sven Büchel

Jena University Language & Information Engineering
(JULIE) Lab

<http://www.julielab.de>

Friedrich-Schiller-Universität Jena

Geisteswissenschaften und Emotionen



https://de.wikipedia.org/wiki/Walther_von_der_Vogelweide

<https://de.wikipedia.org/wiki/Laokoon-Gruppe>

https://de.wikipedia.org/wiki/Der_Schrei

<http://www.br.de/telekolleg/faecher/deutsch/literatur/goethe-weimarer-klassik-100.html>

Unterschiedliche Datendichten

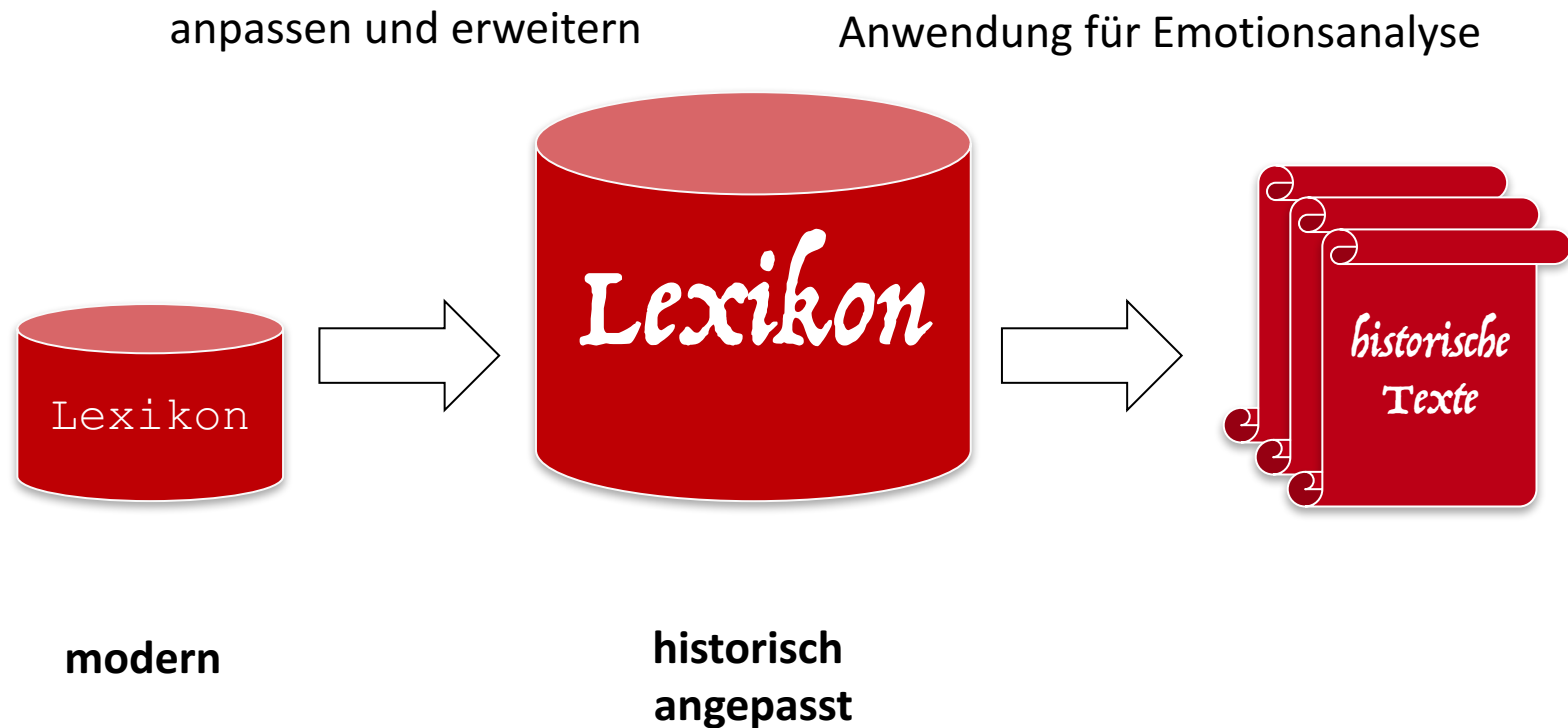
Zeitgenössische Sprachdaten



Historische Sprachdaten

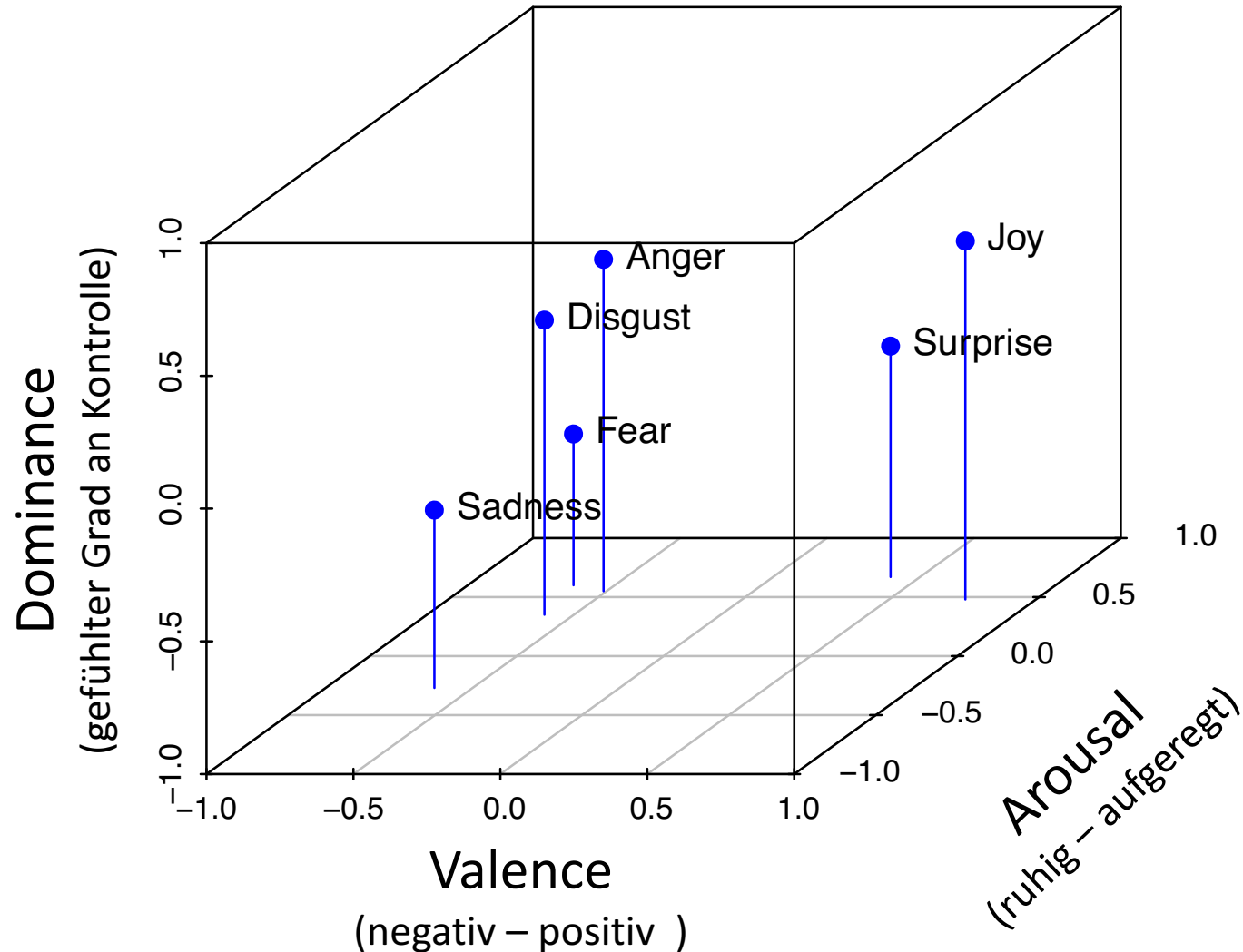


Methodisches Vorgehen



Methoden

Formale Darstellung von Emotionen: VAD



Emotionslexika

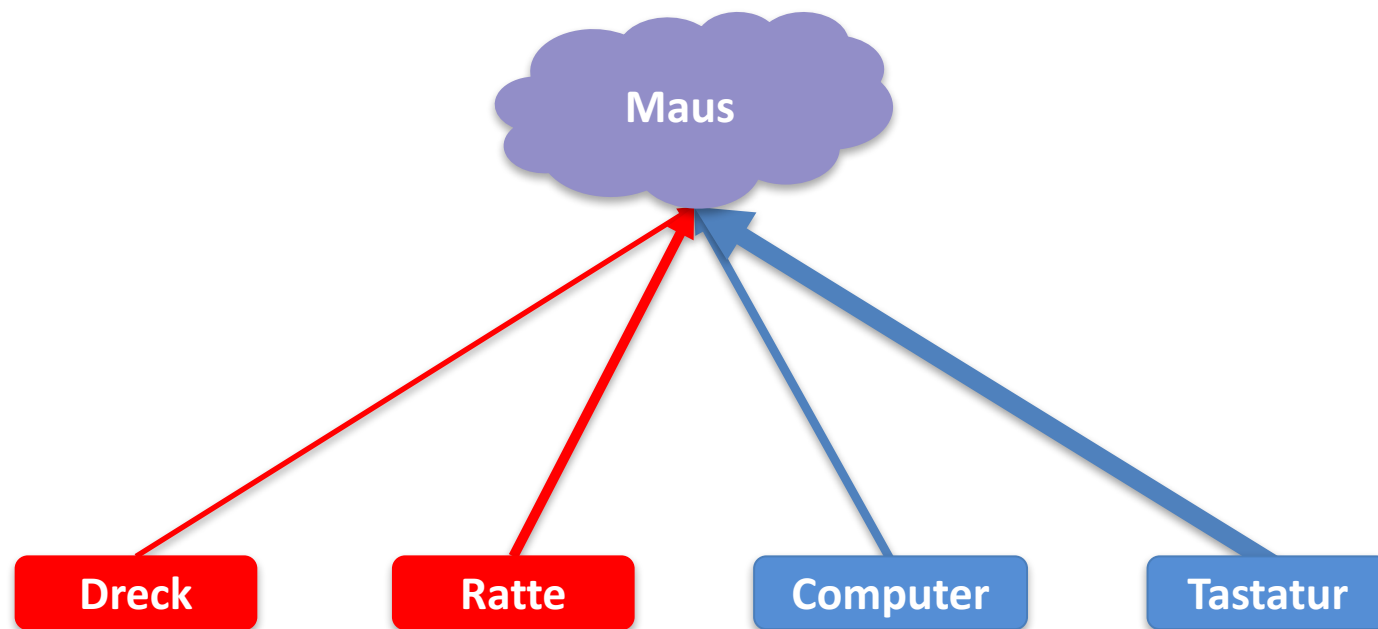
- Lexikalische Datensätze
- Kodiert Emotionen einzelner Wörter
- Kontextunabhängig
- Zahlreiche Datensätze in Psychologie aufgebaut

Lemma	Valence	Arousal	Dominance
Sonnenschein	8.1	5.3	5.4
Terrorismus	1.6	7.4	2.7
Ekstase	6.9	7.8	3.0

(9-Punkt-Skala)

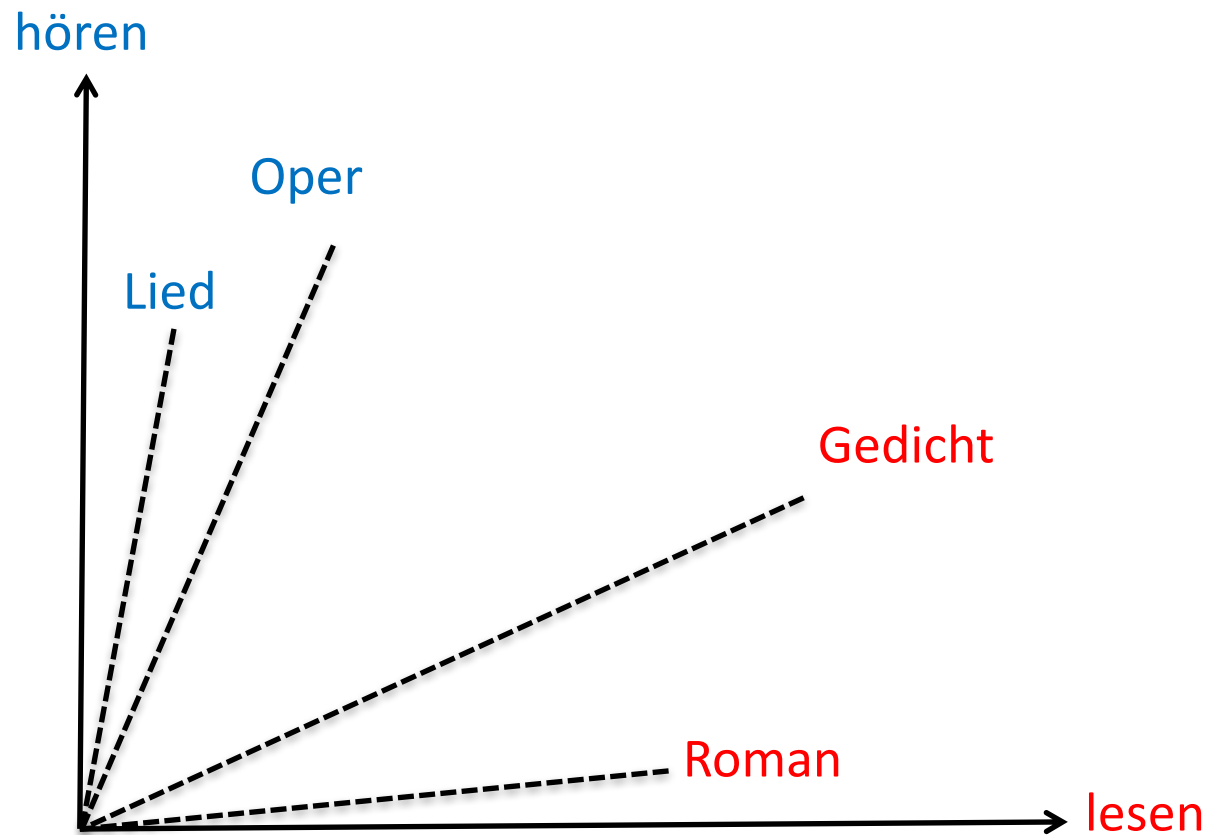
Automatische Erweiterung von Emotionslexika

- Basierend auf sog. *seed*-Wörtern
- Turney-Littman-Algorithmus

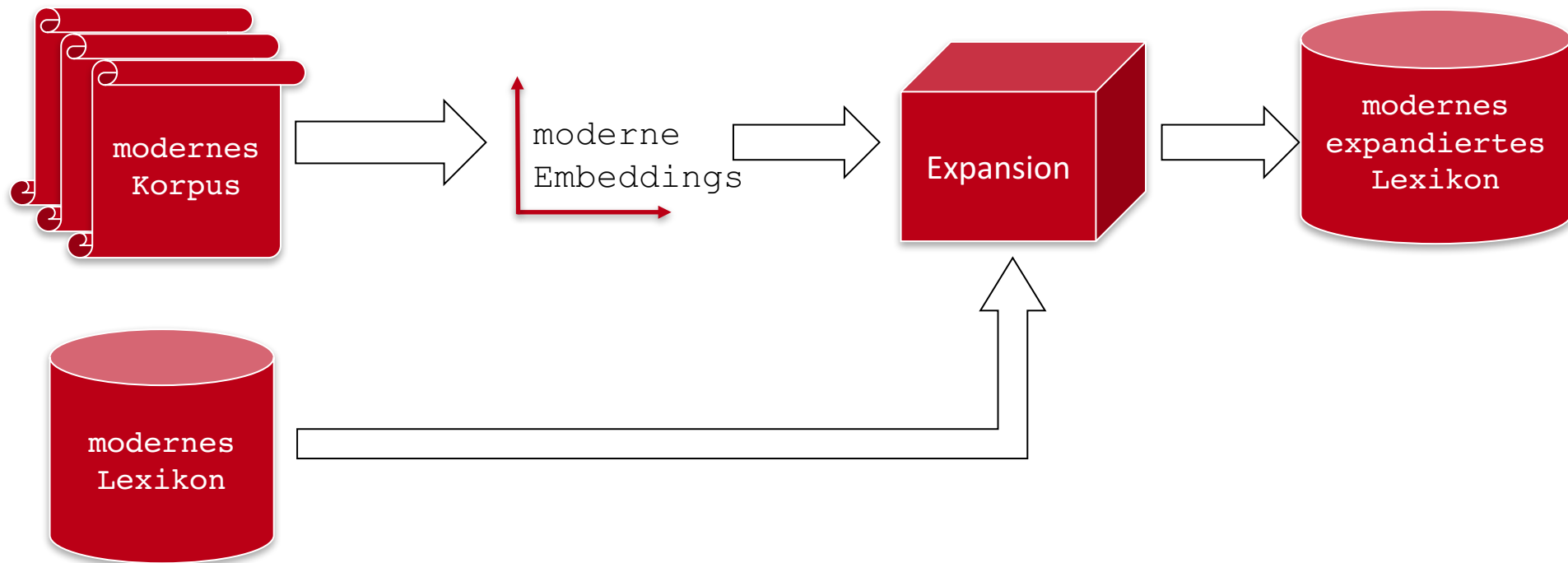


Quantifizierung von Wortähnlichkeit

word embeddings



Synchrone Lexikon-Expansion



Diachrone Lexikon-Expansion

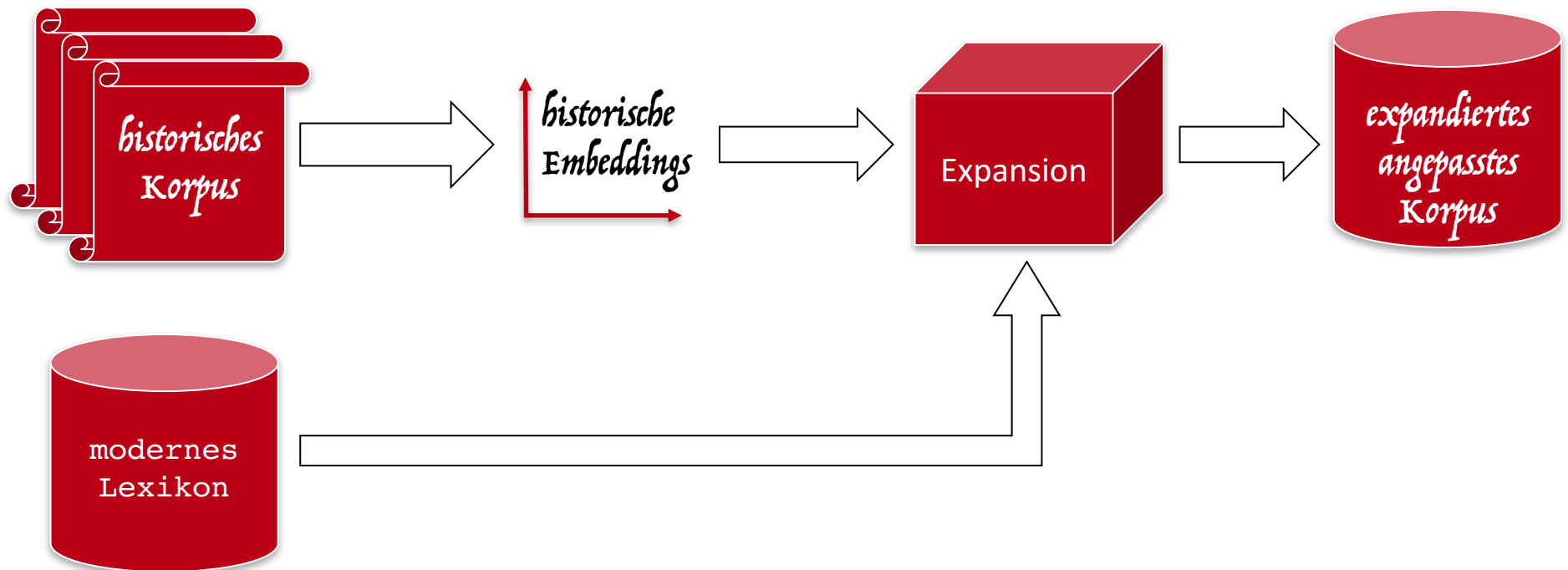


Illustration des Anpassungsschritts (vorher)

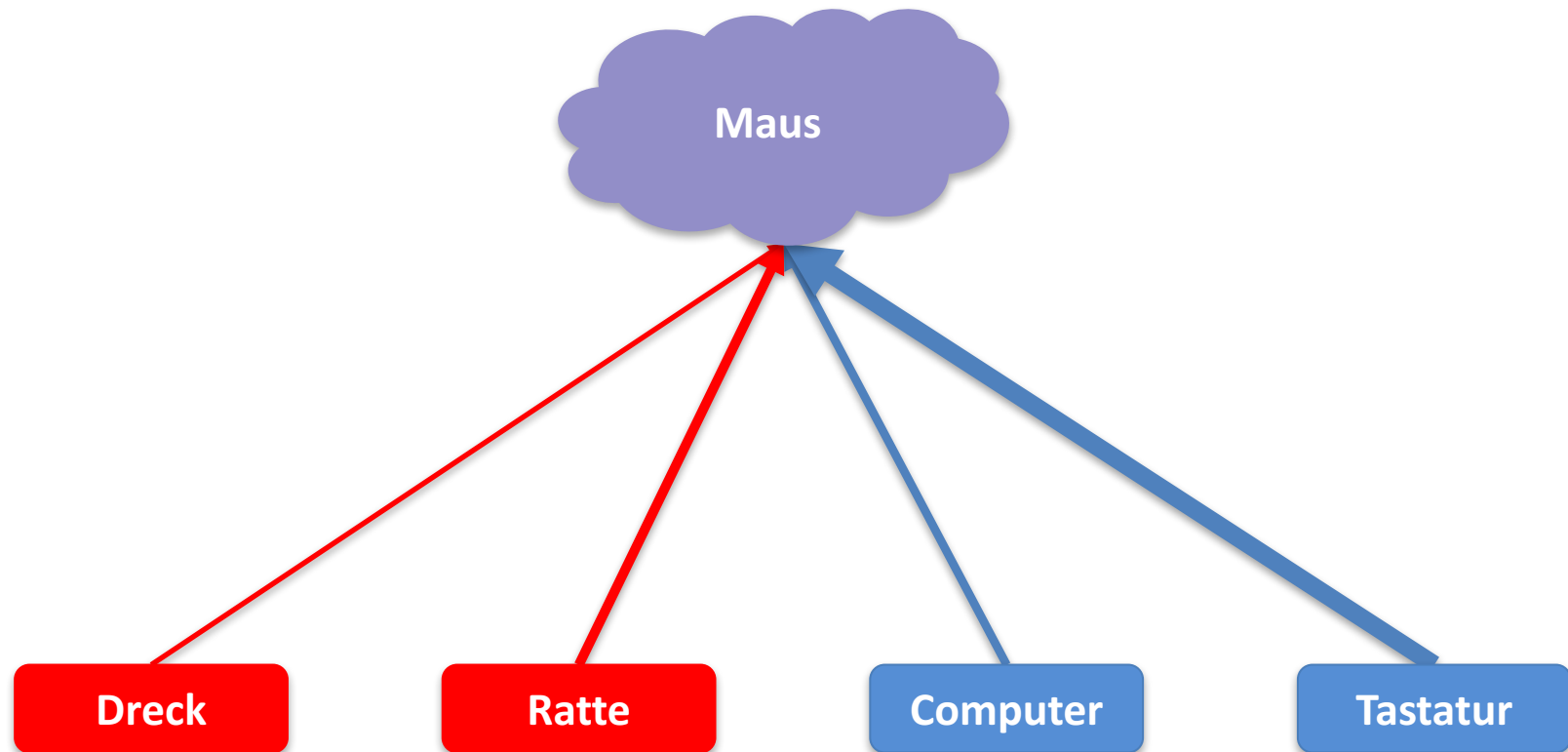
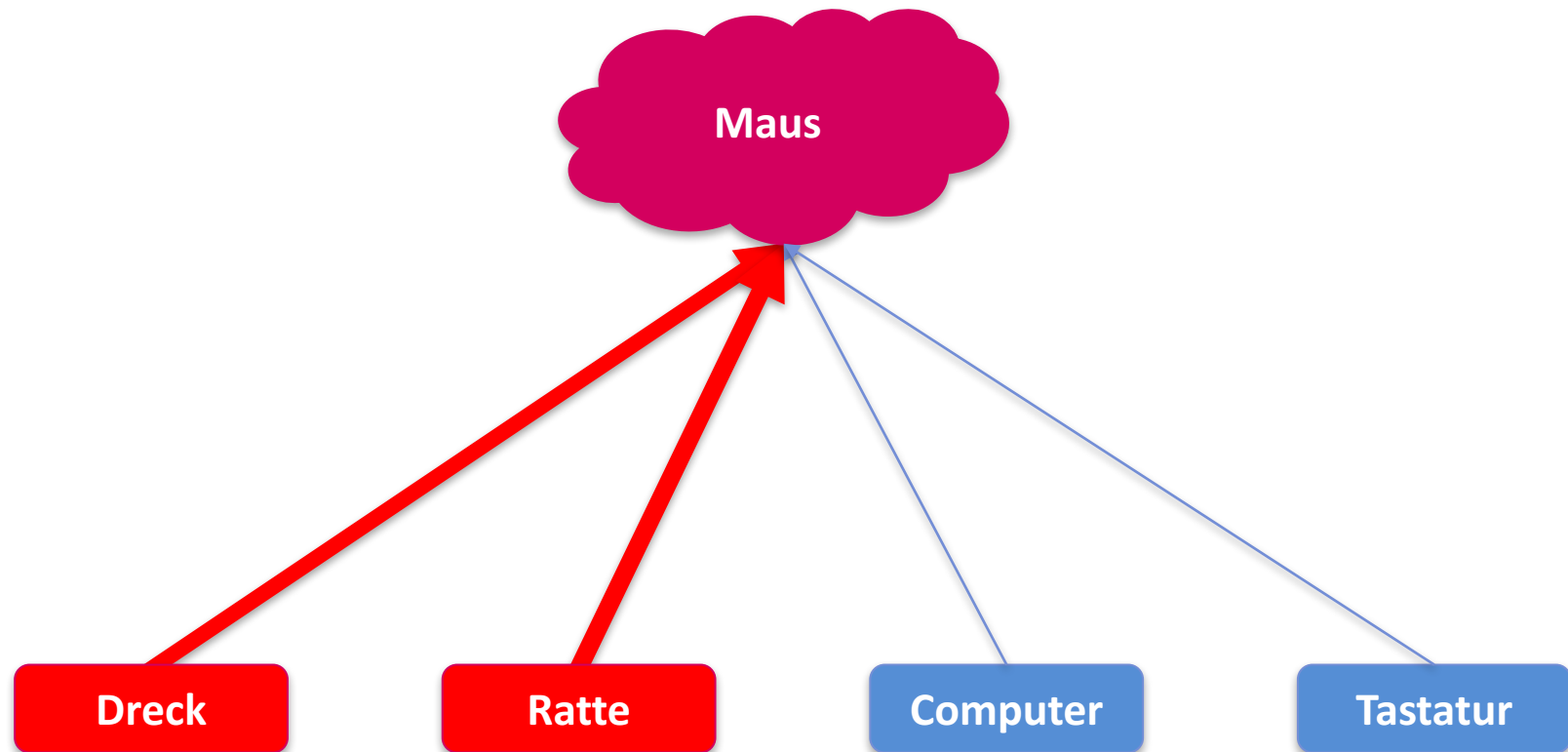


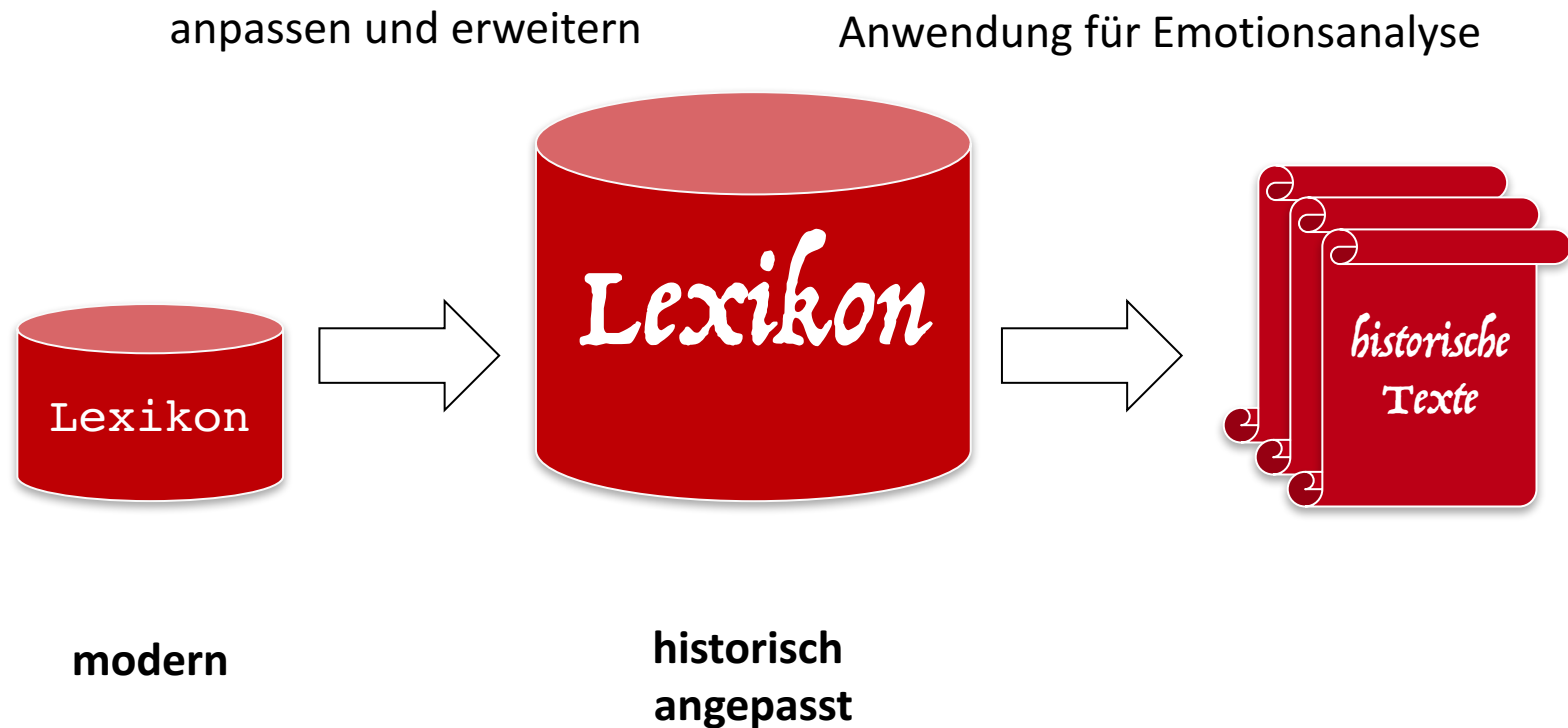
Illustration des Anpassungsschritts (nachher)



Experiment I

(Quantifizierung historischer Wortemotionen)

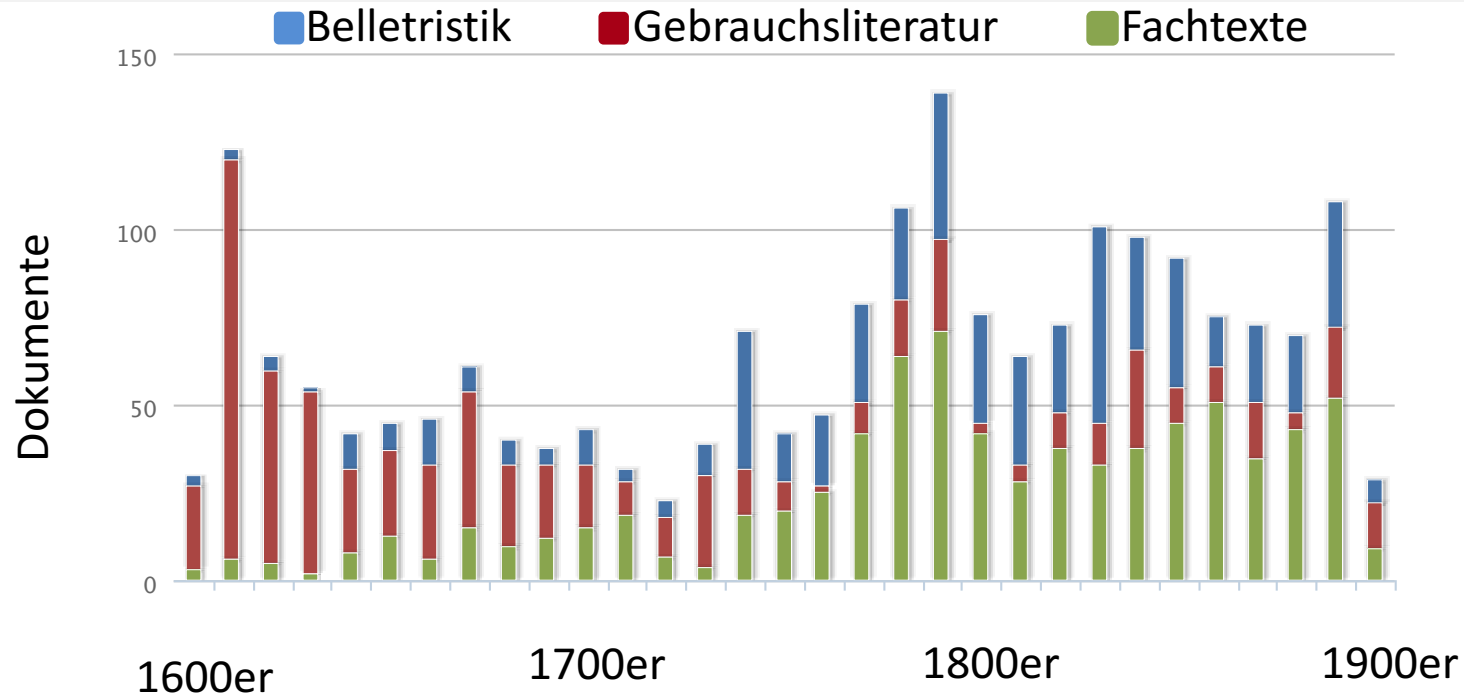
Methodisches Vorgehen



Ressourcen

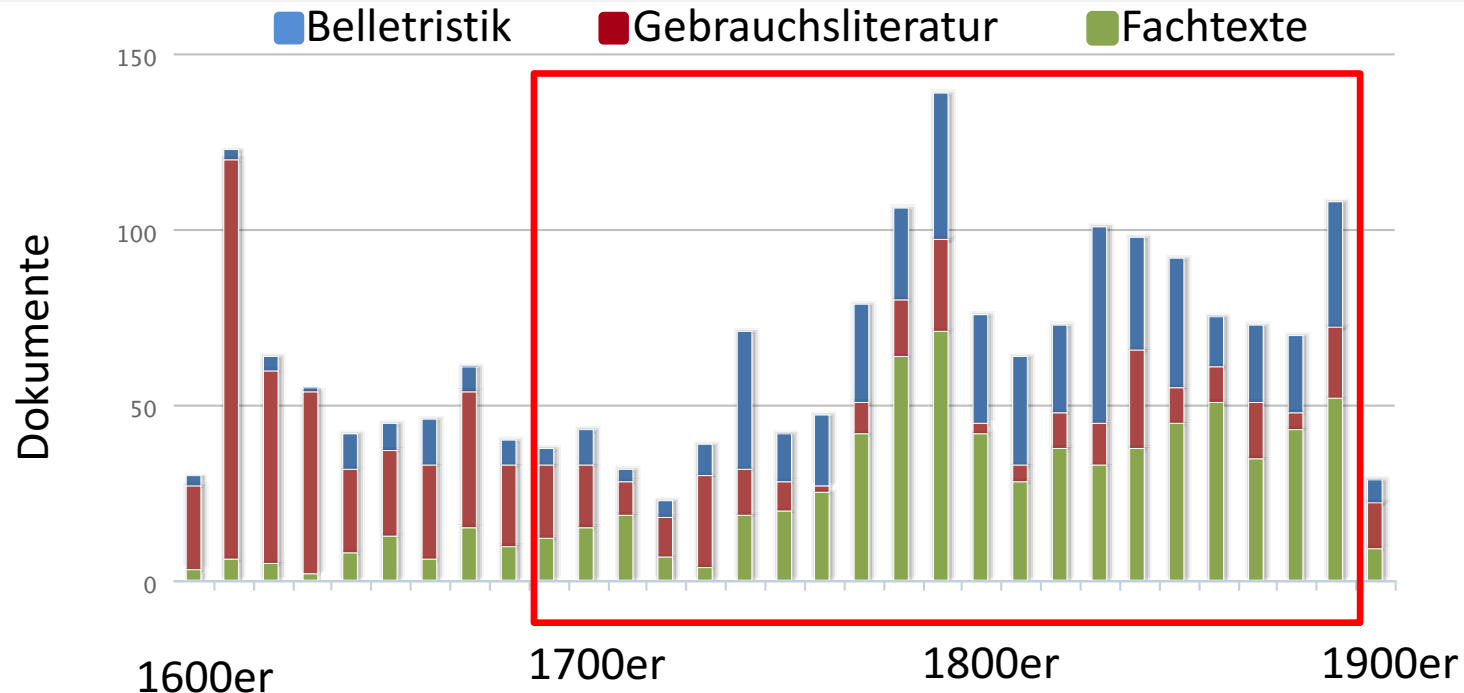
- *Seed*-Lexikon:
 - ANGST (Schmidtke et al. (2014))
 - 1.000 deutsche Wort-Emotionspaare
- Zielkorpus:
 - Deutsches Textarchiv (DTA)
 - viele Meta-Daten verfügbar (z.B. Klassen und Unterklassen)

Zielkorpus: DTA



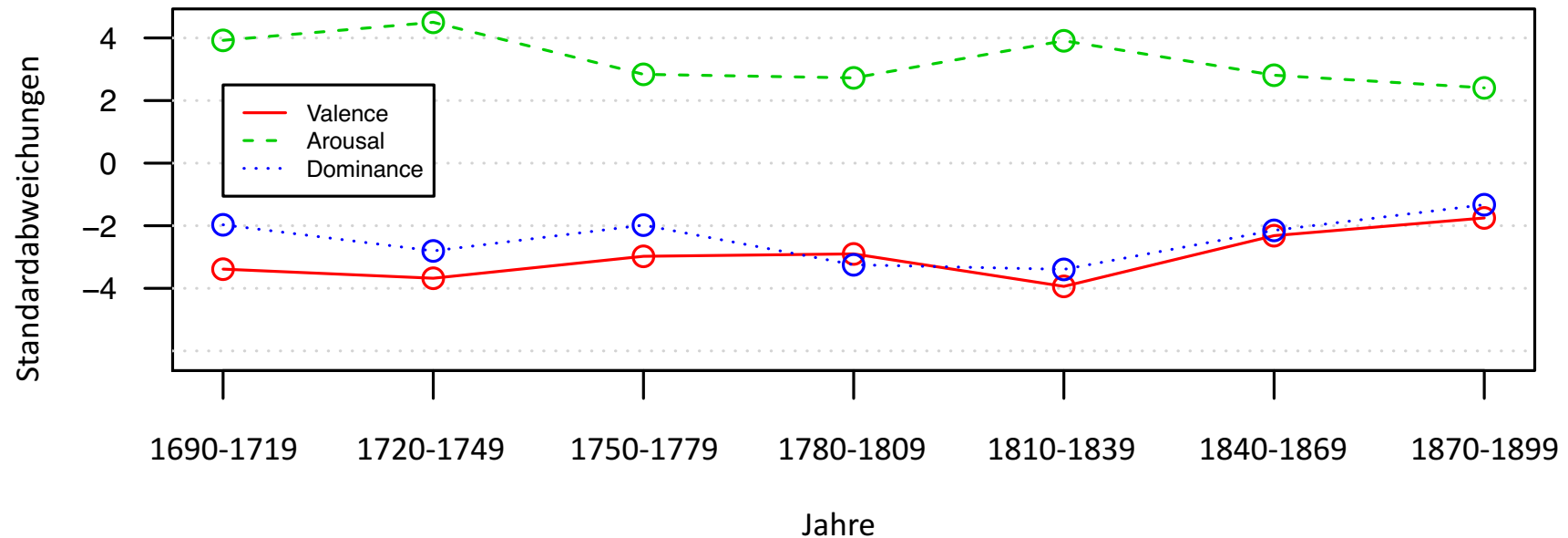
- Aggregation von jeweils 30-Jahres-Abschnitten
- Auswahlzeitraum 1690-1899 (~ 1000 Dokumente, 7 Abschnitte)

Zielkorpus: DTA



- Aggregation von jeweils 30-Jahres-Abschnitten
- Auswahlzeitraum 1690-1899 (~ 1000 Dokumente, 7 Abschnitte)

Emotionsverlauf von *Sünde*



- Starker Anstieg der Valenz seit Beginn der Aufklärung
- Vergleich mit korpuslinguistischen Methoden

Die 10 häufigsten Kollokationen von *Sünde*

Rang	Lemma	
	1690er	1890er
1	todt-	Lamm
2	Erzürnung	hinwegnehmen
3	läßlich	Verzeihung
4	beichten	Ausschweifung
5	Nachlaß	Gotte
6	Grobheit	Schande
7	verschweigen	Reue
8	beweinen	Ärgernis
9	pichen	Laster
10	beichten	aufrichtig

Die 10 häufigsten Kollokationen von *Sünde*

Rang	Lemma	
	1690er	1890er
1	todt-	Lamm
2	Erzürnung	hinwegnehmen
3	läßlich	Verzeihung
4	beichten	Ausschweifung
5	Nachlaß	Gotte
6	Grobheit	Schande
7	verschweigen	Reue
8	beweinen	Ärgernis
9	pichen	Laster
10	beichten	aufrichtig

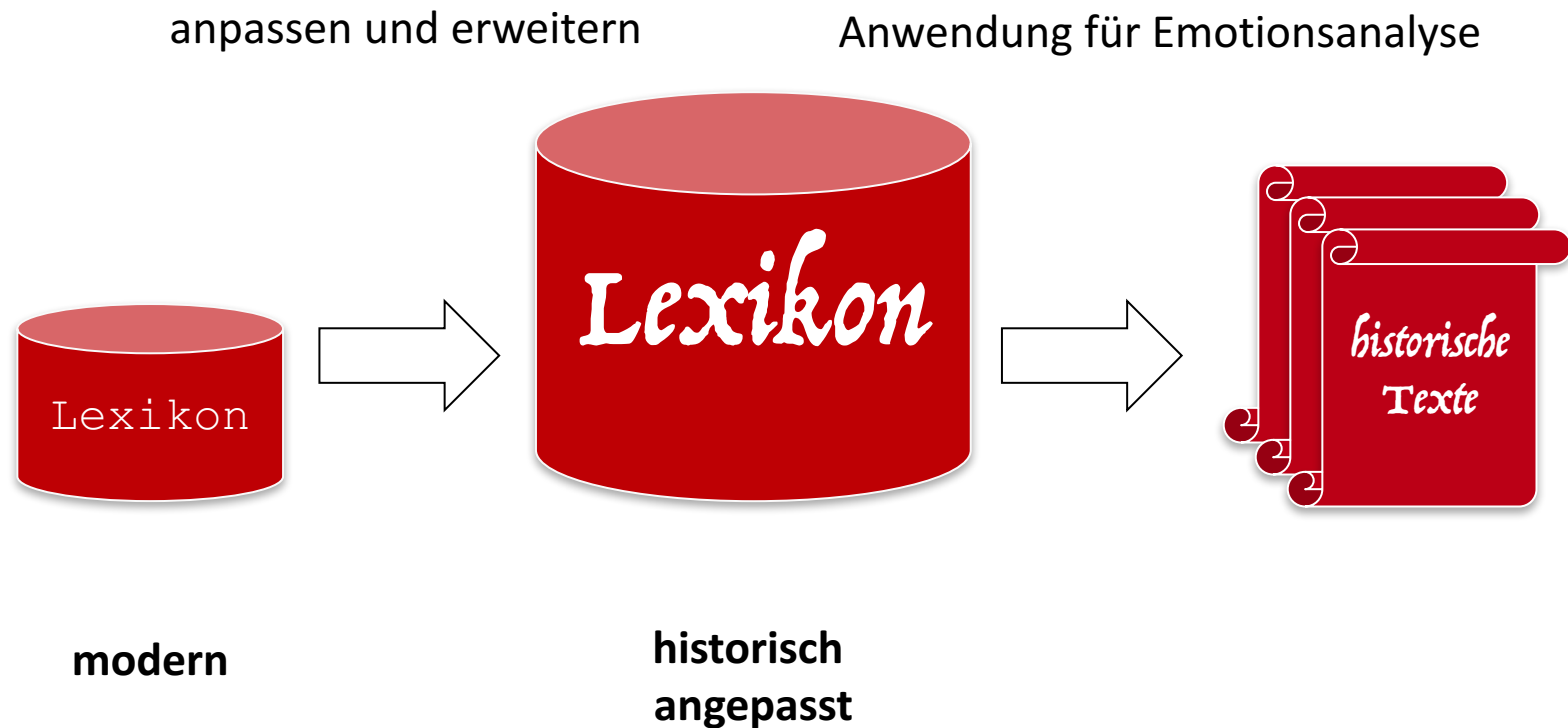
Die 10 häufigsten Kollokationen von *Sünde*

Rang	Lemma	
	1690er	1890er
1	todt-	Lamm
2	Erzürnung	hinwegnehmen
3	läßlich	Verzeihung
4	beichten	Ausschweifung
5	Nachlaß	Gotte
6	Grobheit	Schande
7	verschweigen	Reue
8	beweinen	Ärgernis
9	pichen	Laster
10	beichten	aufrichtig

Experiment II

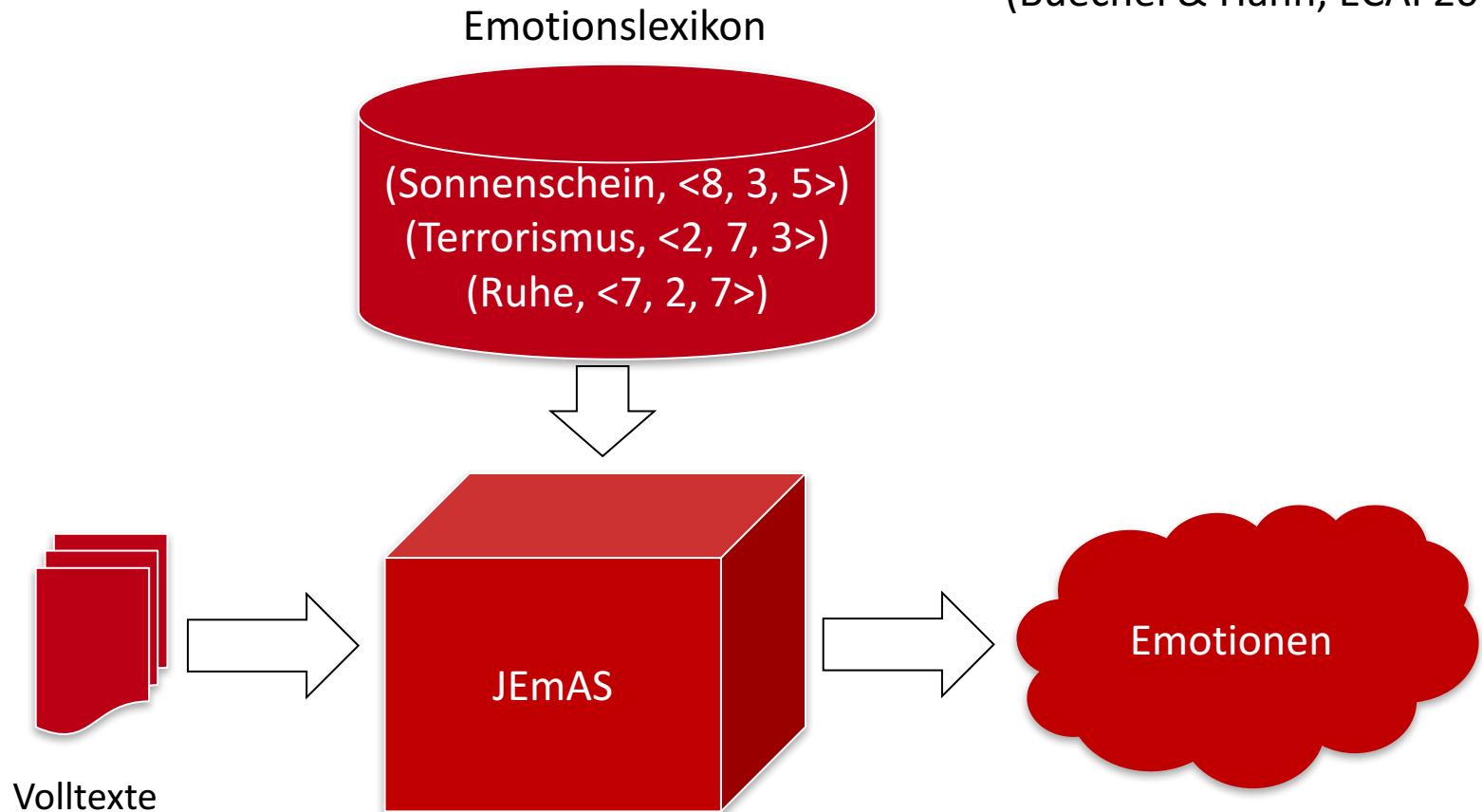
(Quantifizierung historischer Textemotionen)

Methodisches Vorgehen



Textbasierte Emotionsanalyse: JEmAS

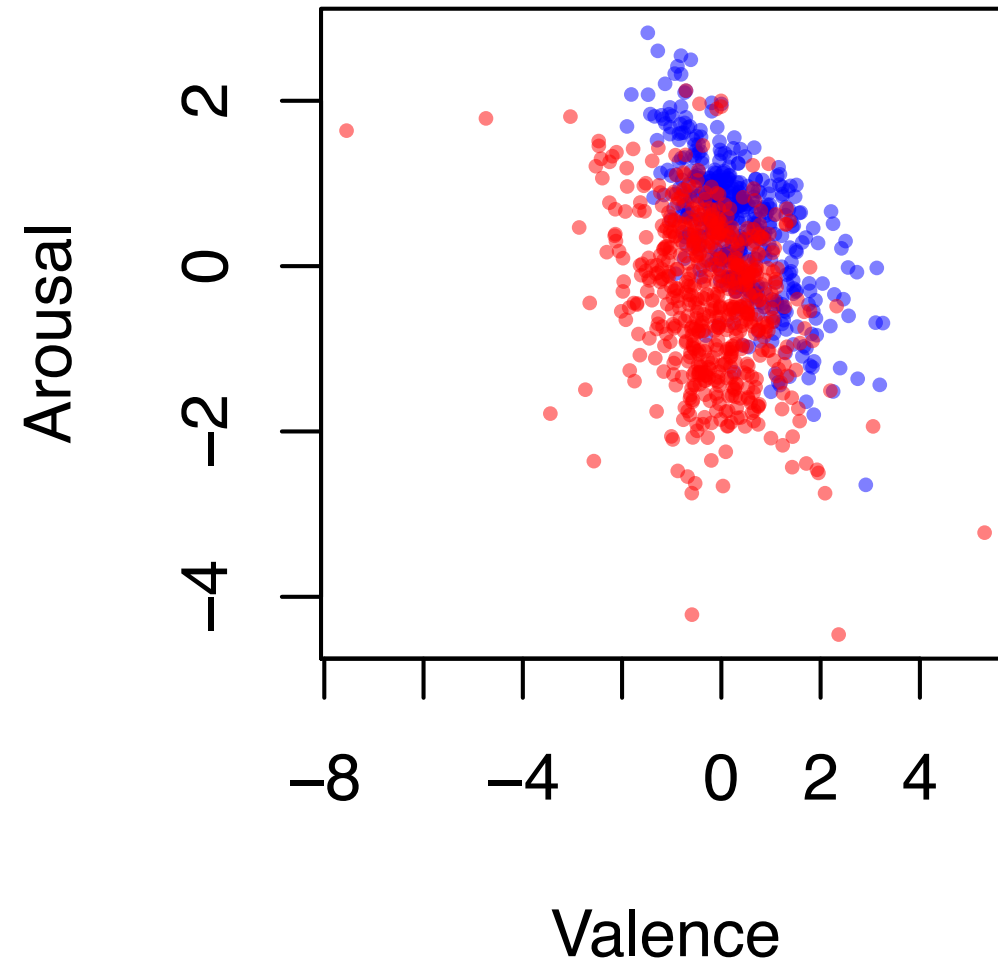
(Buechel & Hahn, ECAI 2016)



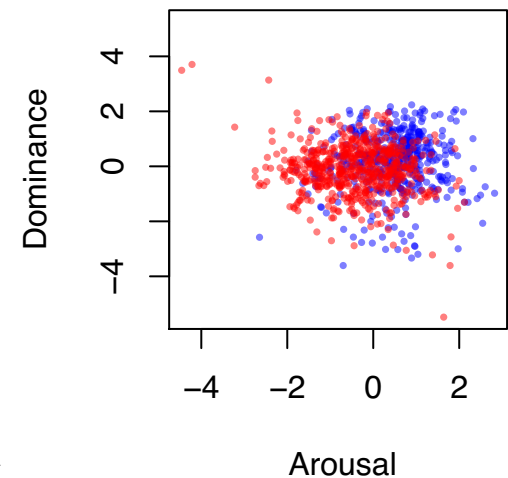
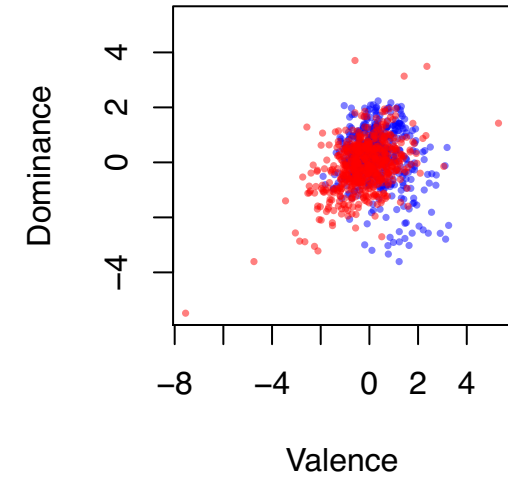
Verfügbar:

<https://github.com/JULIELab/JEmAS>

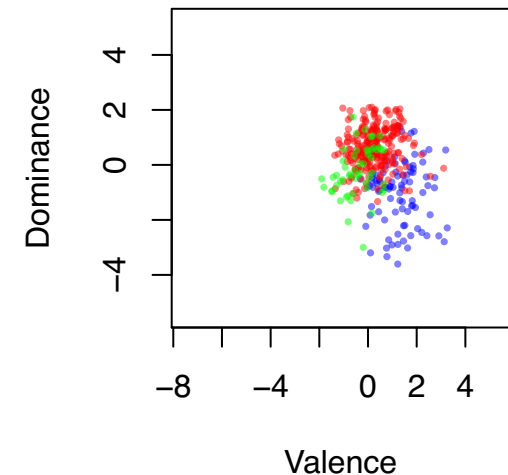
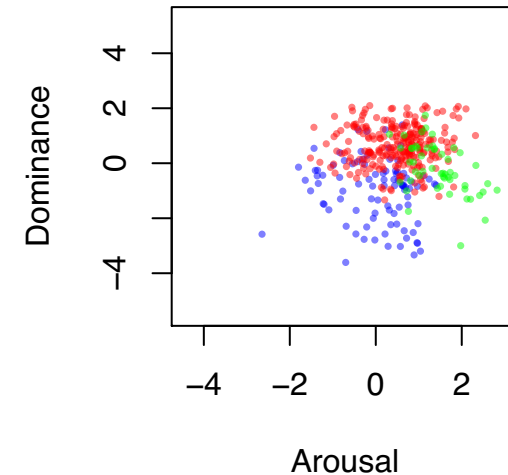
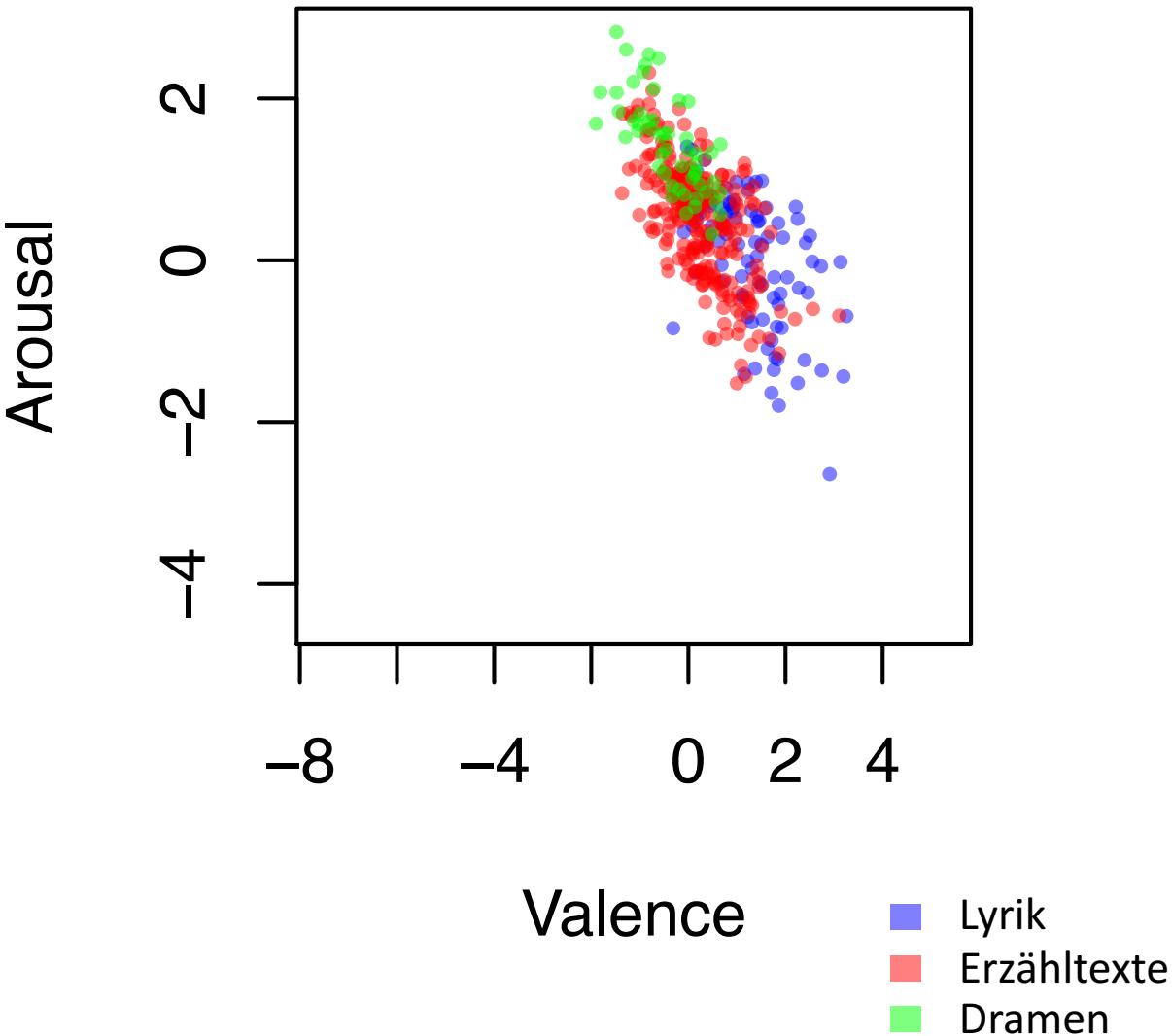
Unterscheidung der DTA-Textsorten



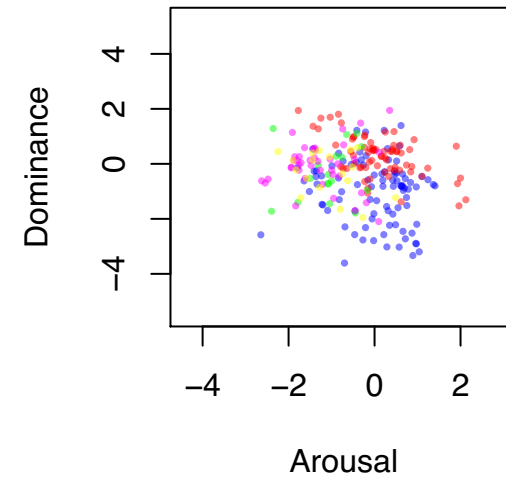
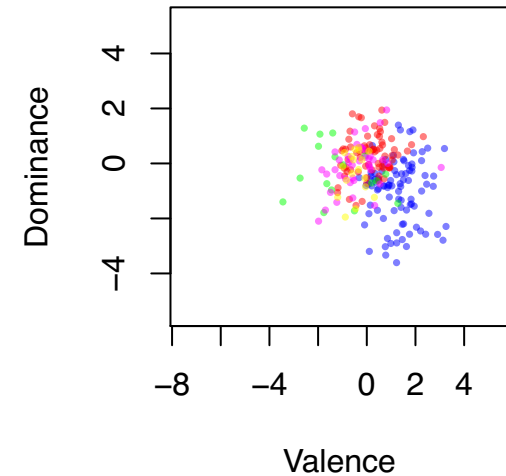
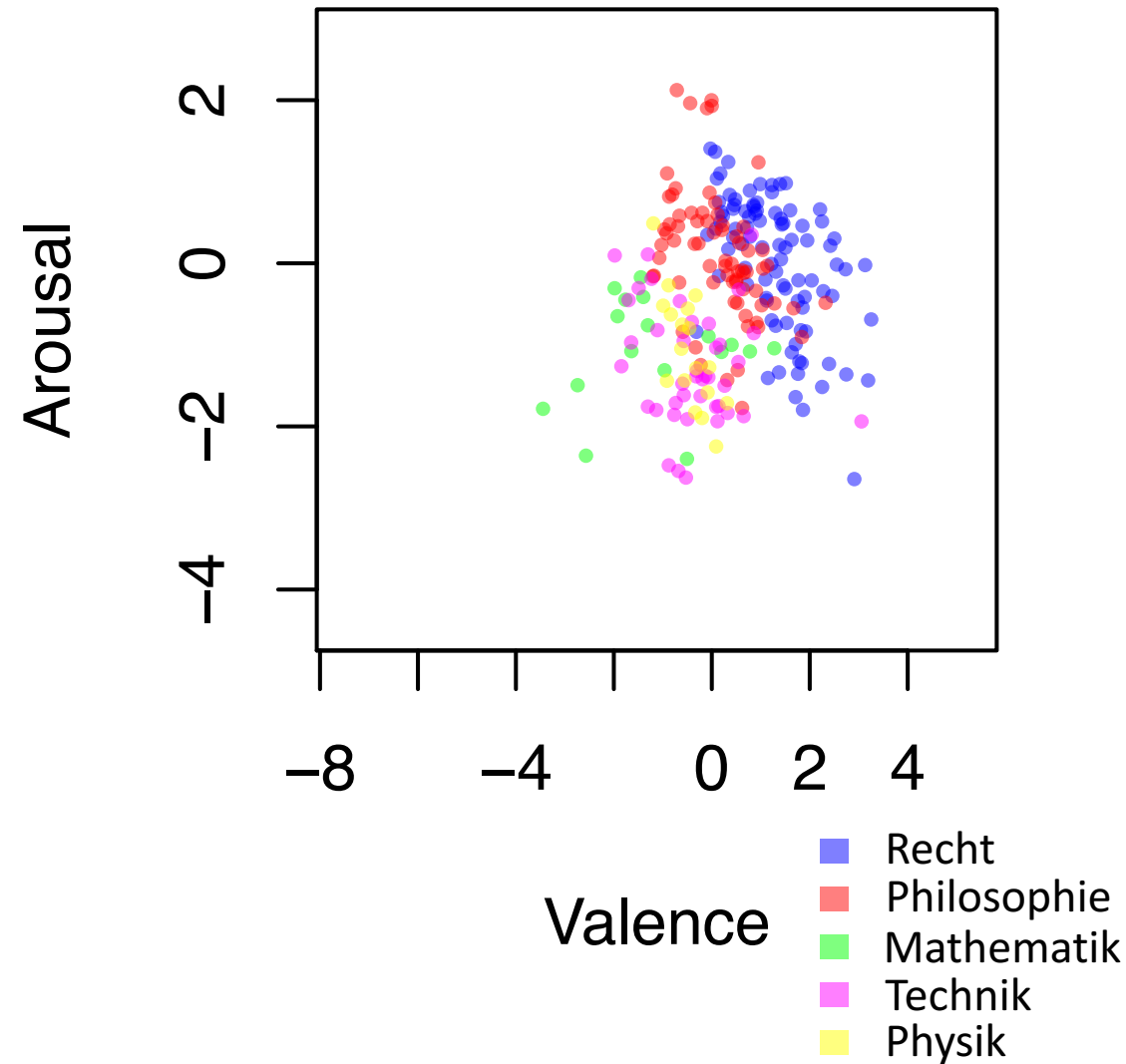
■ Belletristik
■ Fachtexte



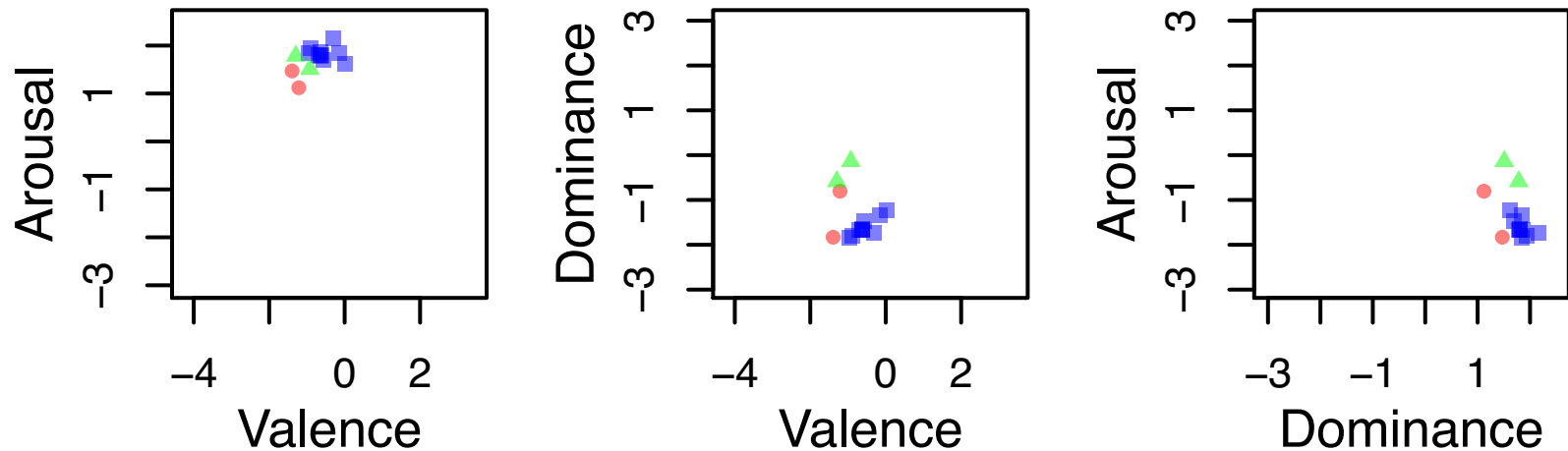
Unterscheidung literarischer Gattungen



Unterscheidung akademischer Disziplinen

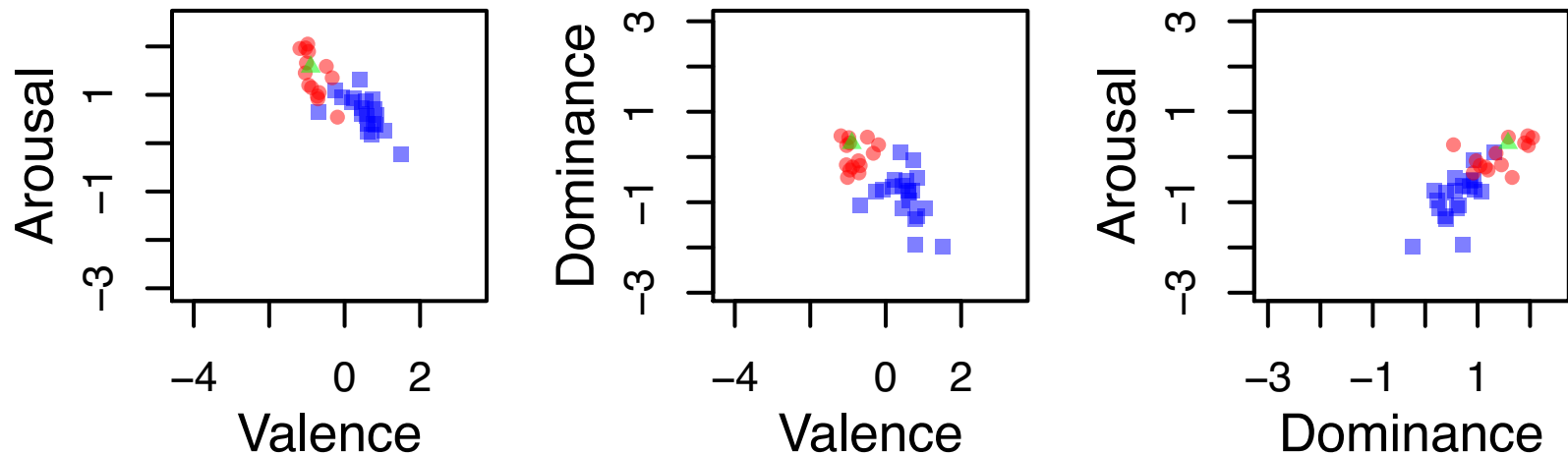


Entwicklung literarischer Gattungen (1690-1719)



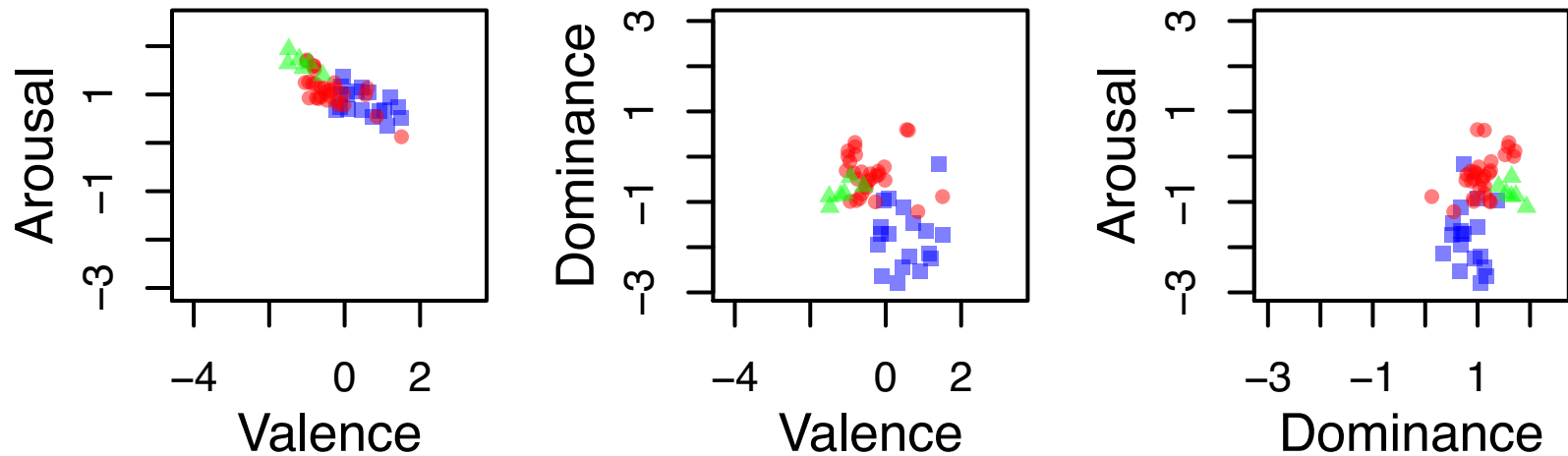
■ Lyrik ● Erzähltexte ▲ Dramen

Entwicklung literarischer Gattungen (1720-1749)



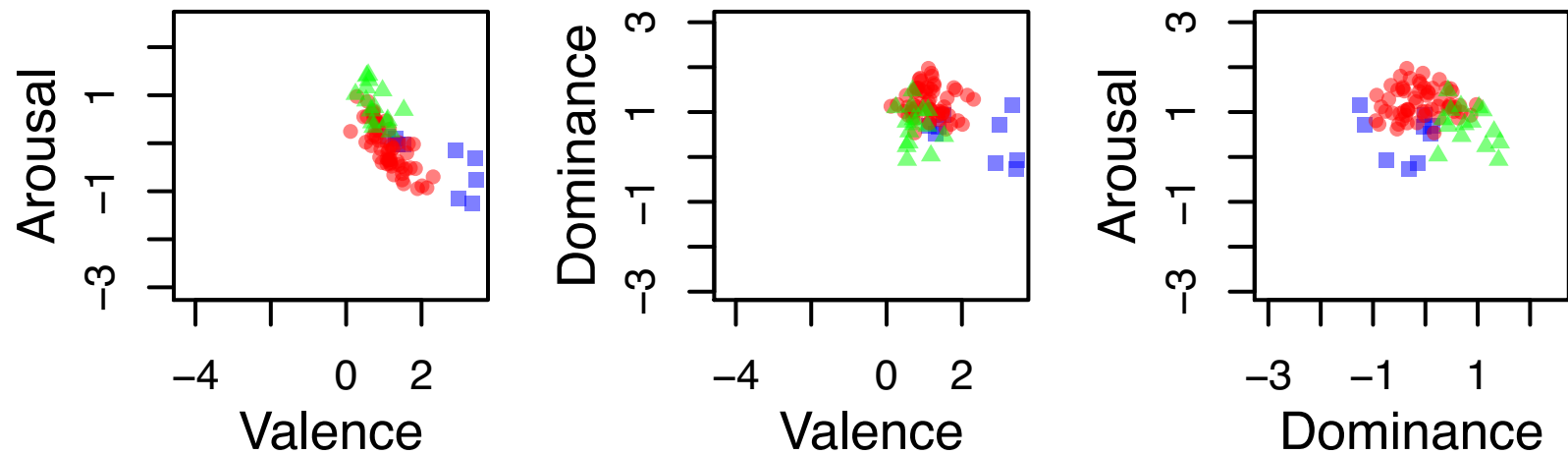
■ Lyrik ● Erzähltexte ▲ Dramen

Entwicklung literarischer Gattungen (1750-1779)



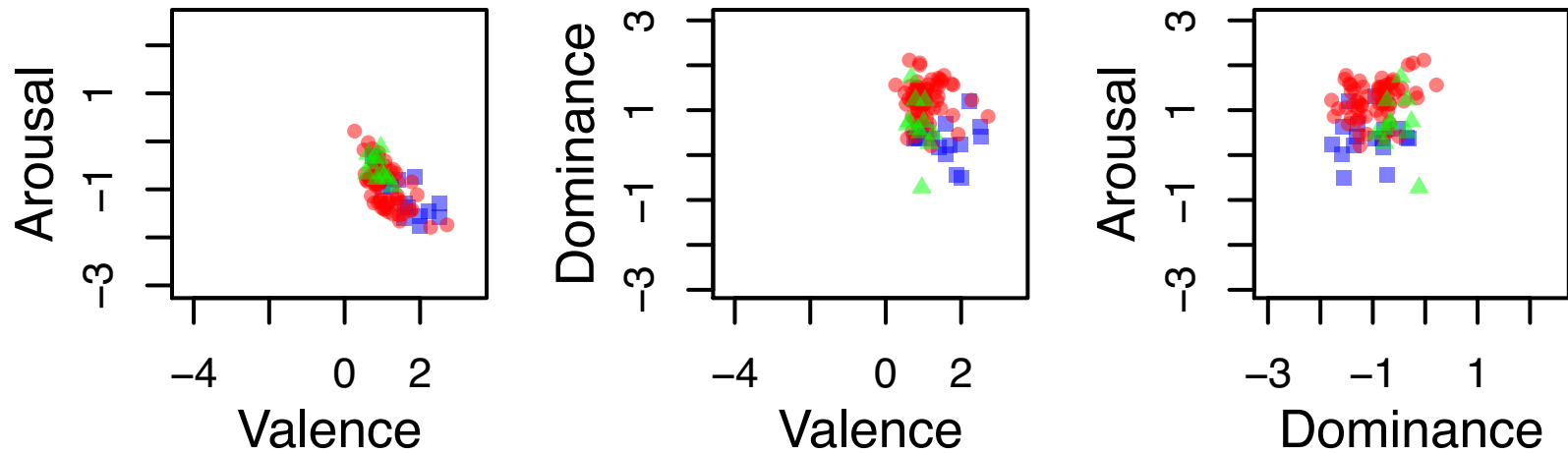
■ Lyrik ● Erzähltexte ▲ Dramen

Entwicklung literarischer Gattungen (1780-1809)



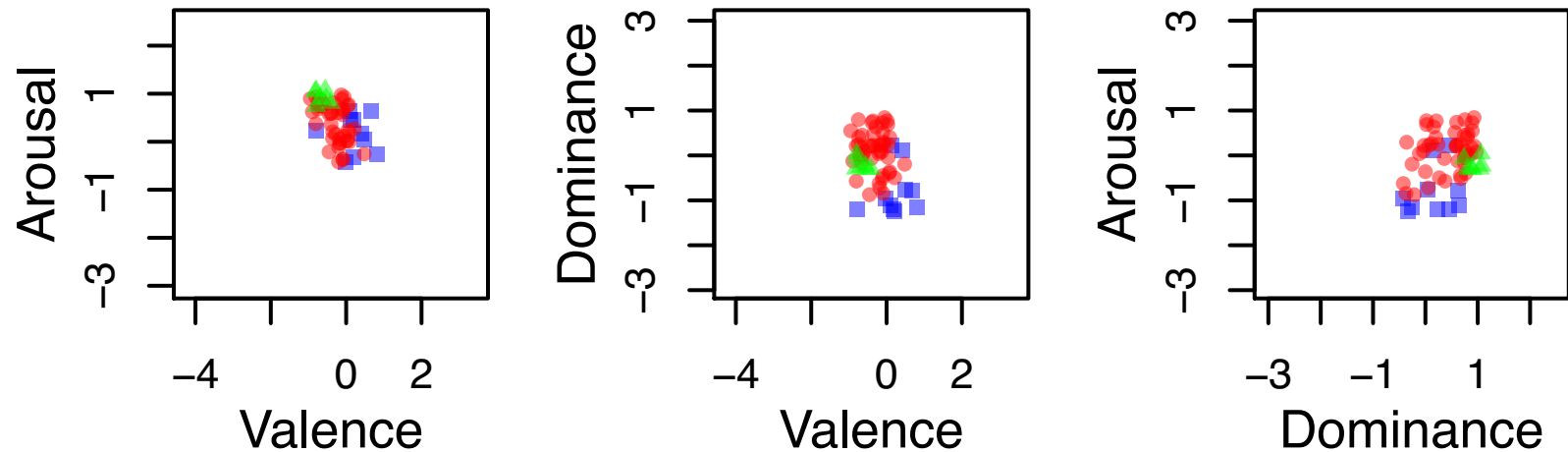
■ Lyrik ● Erzähltexte ▲ Dramen

Entwicklung literarischer Gattungen (1810-1839)



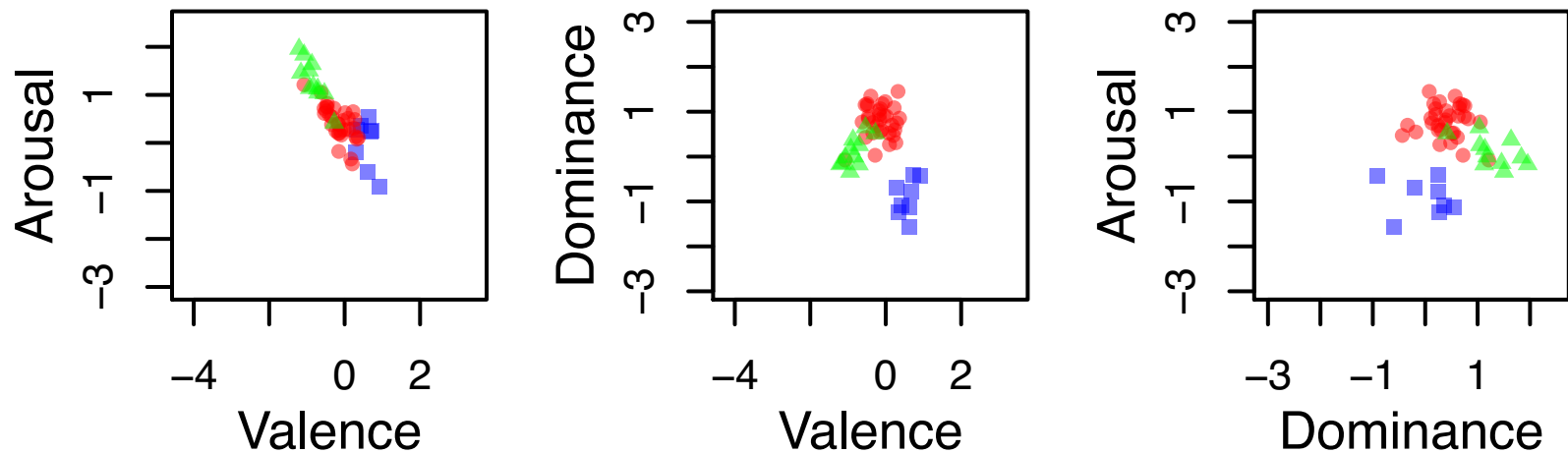
■ Lyrik ● Erzähltexte ▲ Dramen

Entwicklung literarischer Gattungen (1840-1869)



■ Lyrik ● Erzähltexte ▲ Dramen

Entwicklung literarischer Gattungen (1870-1899)



■ Lyrik ● Erzähltexte ▲ Dramen

Zusammenfassung

Zusammenfassung

- Methode für die Quantifizierung von Emotionen in historischer Sprache
- Kombination aus Sprachstufenanpassung für Emotionslexika und lexikonbasierter Textanalyse
- Fallstudie zur Anwendung auf das DTA-Korpus
- Qualitative Validierung aus Wort- und Textebene
- Gattungen zeigen distinktive, zeitlich veränderliche Emotionsprofile
- Ressourcen verfügbar: <https://github.com/JULIELab/HistEmo>
- Webportal zum Abfragen der Emotionswerte: <http://jeseme.org>

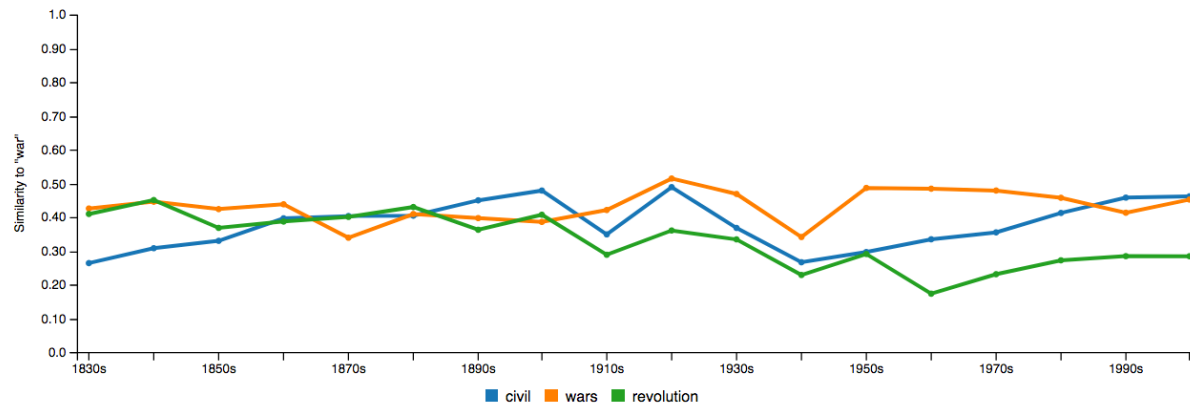
JeSemE 2.0 - The Jena Semantic Explorer

Results for "war" in Corpus of Historical American English

Note: lowercased

Search in [Corpus of Historical American English](#)

Similar Words

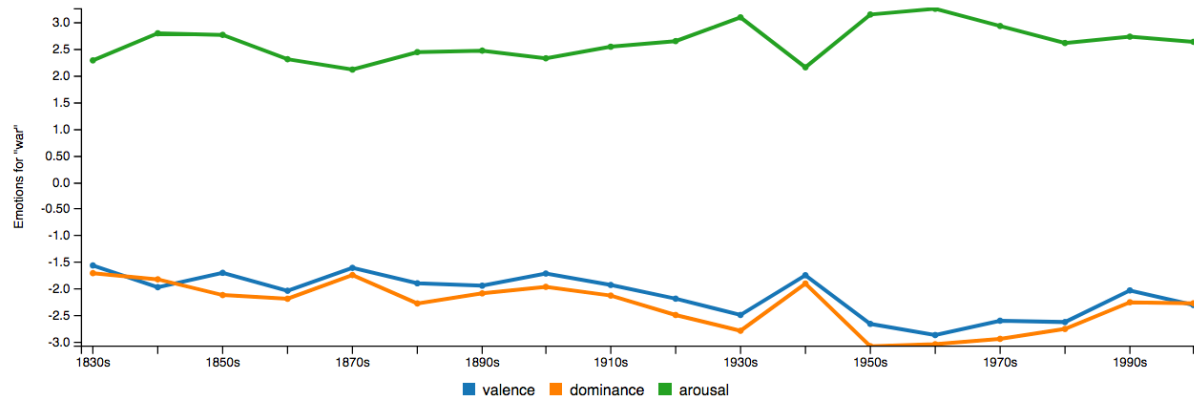


This chart shows the words detected as most similar to "war" and how their similarity changed over time (higher is more similar). These values are based on SVD_{PPMI} — see [Help](#) for details.

Add word to graph:

add

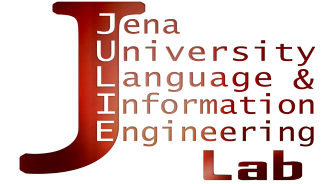
Word Emotion



This chart shows how the emotions connected to "war" changed over time. We describe emotions via valence (positive – negative), arousal (calm – excited) and dominance (controlled – in control) — see [Help](#) for details. Scores are given in standard deviations around an (for each time span) emotionally neutral 0.



FRIEDRICH-SCHILLER-
UNIVERSITÄT
JENA



Quantifizierung von Emotionen in historischer Sprache

Sven Büchel

Jena University Language & Information Engineering
(JULIE) Lab

<http://www.julielab.de>

Friedrich-Schiller-Universität Jena